A TALE OF TWO CITIES: EXAMINING NONPROFIT ACCOUNTABILITY

THROUGH PROGRAM EVALUATION

by

MARCELA NAVA

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN SOCIAL WORK

THE UNIVERSITY OF TEXAS AT ARLINGTON

August 2009

ACKNOWLEDGEMENTS

ABSTRACT

A TALE OF TWO CITIES: EXAMINING NONPROFIT ACCOUNTABILITY

THROUGH PROGRAM EVALUATION

Marcela Nava, MSSW

The University of Texas at Arlington, 2009

Supervising Professor: Richard Hoefer

Nonprofit organizations exist in a world of constantly increasing scrutiny and oversight, while public scandals create an increasing demand for public accountability. Program evaluation is seen as a powerful tool to assist nonprofits in establishing organizational accountability, yet studies suggest that program evaluation is not being used to establish accountability. In order to explore the use of program evaluation in establishing accountability, surveys were administered to nonprofit organizations in Dallas and Fort Worth. Respondents also participated in interviews about decision making processes regarding evaluation and accountability within their agencies.

Results indicate that organizations in Dallas and Fort Worth are fairly similar on many aspects of evaluation and accountability. Key differences in the use of evaluations, role of faith-based agencies, influence of funders, elements of program design, and other aspects of evaluation and accountability are explored.

.

iv

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Nonprofit organizations exist in a climate of constantly growing scrutiny and oversight. The trend toward privatization of services that began in the last quarter of the twentieth century has vastly extended the role of nonprofit organizations through incentives such as service contracts, salaries, and continued existence. As governments increasingly rely on nonprofit organizations to deliver social services, these agencies become involved in meeting more stringent accountability requirements such as program evaluations and performance measures (Lindgren, 2001; Jurik, 2004). Concurrently, widespread concern about the theory and practice of evaluation methods used by nonprofit organizations has led to a growing awareness of accountability in the nonprofit sector.

Recent events such as corporate accountability scandals, abuse and exploitation within religious institutions and the fiscal mismanagement of national nonprofit organizations have also led to public outcry for greater accountability (Kearns, 1996; Walden, 2006; Smith and Richmond, 2007). Nonprofit organizations have experienced the consequences from publicized scandals through a loss of contributions, diminished status, and organizational turmoil. As a result, a greater push for increased accountability has come about through the combined efforts of funding sources, donors, and the general public (Gibelman and Gelman, 2000).

This article expands on Hoefer's research (2000, 2008) examining changes in agency use of evaluation to ensure maximum accountability. Utilizing original data from Hoefer's research in Dallas, Texas, as well as newly collected data from the Fort Worth, Texas area, this study compares nonprofit agencies in two North Texas regions, Dallas County and Tarrant County, to better understand the use of program evaluation, the use of evaluation to promote

1

accountability, attitudes about evaluation, and the decision-making processes related to evaluation and accountability within the agencies.

CHAPTER 2

LITERATURE REVIEW

The concept of accountability in the nonprofit sector typically refers to the practice of performance measurement, reporting, and evaluation. However, the term is often used so generically as to undermine the construct's utility. Cutt and Murray (2000) provide a general definition of accountability as "shared expectations expressed through a common currency" (p.1). Accountability exists on several planes such as the political, market, and administrative fields and can focus on finances, outcomes, or processes (Brown and Troutt, 2007). Cutt and Murray concur that all forms of accountability include the following key elements: the obligation to deliver an account for a responsibility that has been conferred; the existence of two parties, one who allocates responsibility and one who accepts it; and the formal requirement of a disclosure through some specified mechanism (2000).

Nonprofit organizations in the social sector face unique challenges in achieving accountability. As agencies serving a public purpose, nonprofits must demonstrate the impact of their programs, while their legal status demands evidence of the appropriate use of their resources. Nonprofit agencies are therefore held to standards of effectiveness (relating program activities to outcomes) and efficiency (relating costs to outcomes). Traditional notions of accountability focus on procedural accountability, which reflects the use of fiduciary disclosure to demonstrate appropriate stewardship of resources (Cutt and Ritter, 1984). This type of accountability requires that the party accepting responsibility correctly follow procedures set forth by the party allocating responsibility (Brown and Troutt, 2007). Performance (or consequential) accountability, on the other hand, creates mutual accountability between both parties and to the people they serve through outcomes achieved. This type of accountability includes information on the outputs or benefits provided in relation to the inputs (Brown and Troutt, 2007; Cutt and Ritter, 1984).

The increasing focus on accountability has been met with both enthusiasm and resistance by nonprofit agencies. Researchers have borrowed from the medical field to develop evidence-based practice (EBP), an approach that involves using research-based knowledge to guide practitioners, as a method of improving public confidence in the social work profession and applying funds to scientifically tested interventions (Witkin and Harrison, 2001; Hall, 2008). Concurrently, however, nonprofits have encountered barriers to the implementation of evidence-based models (Giogia and Dziadosz, 2008), while research has suggested that increasing performance accountability may hinder the relational work of nonprofits (Benjamin, 2008).

Both program evaluation and basic research engage in a similar investigative process using systematic problem-solving techniques. However, program evaluation is usually only applicable to a specific program, whereas basic or theoretical research is conducted to produce generalizable knowledge. The benefits of conducting program evaluation are numerous, and may include helping a program become more effective, creating a more integrated workforce, and improving both worker and client satisfaction (Hoefer, 2008).

Theorists tend to view program evaluation as either a rational process or a political process (Tassie, Murray and Cutt, 1998). Viewed as a rational process, program evaluation is a tool embraced by agencies and funders as a means of pursuing a specific end, in this case accountability to stakeholders such as funders or the general community. Program evaluation can thus be defined as "applied research used as part of the managerial process" (Royse, Thyer, Padgett and Logan, 2006, p.11). A political perspective, however, holds that program evaluations are also conducted for a variety of pragmatic reasons, such as to compete for scarce funds or to meet a funder's requirement. Recognizing that there is much that cannot be measured through objective research, evaluation becomes a process through which different parties strive to create a dominant perspective, with the most powerful side eventually imposing its subjective views (Tassie, Murray and Cutt, 1998).

The literature is replete with texts on how to conduct program evaluation (for example, see Royse, Thyer, Padgett and Logan, 2006; Fitzpatrick, Sanders and Worthen, 2003; Wholey, Hatry and Newcomer, 2004). In practice, however, nonprofit agencies vary widely in their use of program evaluation including different purposes, audiences, evaluators, and data collection methods and various other variables. Additionally, the emergence of participatory action research, with its focus on engaging staff, clients, and other stakeholders in evaluation design and implementation, has challenged traditional models of evaluation while increasing the likelihood that evaluations will be used to improve programs (Fine, Thayer and Coglan, 2000; Dartington, 1998).

More recently, increases in the number of nonprofit organizations and decreases in the amount of corporate contributions have led fiscal intermediaries such as the United Way to expand their roles as economic regulators and community problem solvers ("Editorial," 1982; Brilliant and Young, 2004). Less government funding and increasing donor choice have also led nonprofits to benefit from charitable mutual fund models as websites are increasingly redesigned with innovative fundraising tools (Hoefer, 2008). With the continued privatization of social services, government funding has been found to significantly alter the fiscal spending pattern of nonprofits, suggesting that public funding may affect a nonprofit agency's ability to provide cost-effective services (Frumkin and Kim, 2002).

Knowledge of program evaluation has been widely recognized as a critical factor in an organization's success (Bailey, 2005), yet studies show that social work research lags behind in this area. A study of thirteen major social work journals published over a five-year period concluded that only 3% out of nearly 2,000 articles evaluated programs in a manner that could guide practitioners toward implementing the intervention studied (Rosen, Proctor, and Staudt, 1999). Not surprisingly, research has also shown that social workers tend to use conceptual rationales in lieu of client-centered, research-based, practice or value rationales when making practice decisions (Rosen, Proctor, Morrow-Howell and Staudt, 1995).

5

Overall, most professionals recognize that program evaluation is a critical element for establishing nonprofit accountability. Concurrently, academics acknowledge that social work research lacks the stringency required to effectively guide agencies in the utilization of program evaluations. In fact, few empirical studies have been produced that evaluate the use of program evaluation within agency settings (for example, see Fine, Thayer and Colghan, 2000). The medical field is far more advanced in its research, even having developed a standardized measure to evaluate practitioners' use and attitudes about evidence-based practice (Leach and Gilham, 2008). Existing studies on the use of program evaluation in nonprofit agencies suggest that attitudes about evaluation have improved over the last few years, with 85% of participants in one study citing evaluations as useful (Hoefer, 1999; Hoefer, 2008; Fine, Thayer and Colghan, 2000). However, little is known about the decision-making processes relating to evaluations and accountability within a nonprofit agency. Recent research also indicates progress is being made in the use of evaluation by human service agencies in one metropolitan area, yet no study has examined differences in the use of program evaluation between distinct networks of social service agencies (Hoefer, 2008).

In an effort to produce a framework to guide empirical research into the use of program evaluation, Hoefer (2008) laid out four criteria to determine if evaluation is being used to assure organizational accountability:

(1) Is evaluation being done?

(2) Are appropriately rigorous research methods being used?

(3) Are front-line staff members learning the results of the evaluation? And

(4) Are other stakeholders (including the public) learning the results of the evaluation? (p.3)

Using these criteria, Hoefer concluded that program evaluations conducted by nonprofit social service agencies required improvement in meeting accountability, yet observed that few empirical studies were tracking this process.

6

Nonprofit agencies in the Dallas area and Fort Worth area face unique challenges in accomplishing program evaluation and accountability. Recent scandals in the Fort Worth area include the burial of condemning information in a large county hospital district, the resignation of an urban school district superintendent on malfeasance charges, and the controversial investment of millions of dollars into a sporting facility by a city government. The Dallas area has been faced with widespread financial mismanagement in their largest school district and widespread extortion charges involving various elected officials and city employees. At the same time, fundraising efforts have fallen short at local agencies and United Ways, creating concerns of continuance and sustainability.

Despite the similar issues faced by both the Dallas and Fort Worth areas, it is unknown how their regional differences may impact the use of program evaluation within nonprofit agencies. For example, close to a third of children in Dallas are living in poverty, while less than a quarter of children in Fort Worth live in similar conditions (Annie E. Casey Foundation, 2007). With a population approximately double that of Fort Worth, Dallas faces more issues related to urbanization than its neighbor 30 miles to the West. Popularly, Fort Worth is considered to be "where the West begins" and maintains strong cultural ties to its history in the cattle trade. Dallas, on the other hand, is seen as a cosmopolitan hub more closely in line with the metropolises of the East. While racial clashes in Dallas have been highly publicized, similar tensions in Fort Worth appear to take place on a more suppressed level. Comparing the Dallas and Fort Worth areas in regards to program evaluation could yield insights into the cultural and demographic factors that affect the use of evaluation.

This study proposes to answer the following empirical questions:

1. How do nonprofit agencies in the Dallas and Fort Worth areas differ in terms of conducting program evaluations, types of evaluations conducted, use of design and data sources, and purpose of evaluations?

2. How do nonprofit agencies in the Dallas and Fort Worth areas differ in their use of program evaluation and other accountability strategies, depending on the source of funding?

3. How do nonprofit agencies in the Dallas and Fort Worth areas differ in their attitudes about program evaluation?

4. How do leaders of nonprofit agencies in the Dallas and Fort Worth areas differ in their decision-making process relating to evaluation and accountability within their agencies?

CHAPTER 3

METHODS

This study utilized survey research techniques to gather data from a convenience sample of human service agencies in the greater Dallas and greater Fort Worth areas. Rather than relying on geographical or demographic data to select nonprofit agencies for inclusion in this research, this study depended on established networks of social service agencies that define themselves in these terms. The survey data utilized in this study was collected through both surveys administered in 2009 to Fort Worth nonprofits, as well as data from Hoefer's 2008 study of Dallas nonprofits. Hoefer's 2008 study utilized the "Blue Book" of agencies compiled by the Community Council of Dallas. This list was initially created by Hoefer based on agencies providing direct services to individuals. A systematic process removed agencies providing informational or medical services. Agencies relying primarily on public funds (e.g. school districts, private contractors) were not included in this study, although agencies receiving a percentage of their funding from government streams were part of the sampling frame. Agencies In all, 67 agencies were included in the final list of agencies for the greater Dallas area. After two mailed contacts, Hoefer received a response rate of 49%, with an n=33. Of all responses received, 11 agencies (33%) indicated a willingness to be interviewed.

This researcher identified agencies for the greater Fort Worth area through the 2007 edition of the Blue Book put together by the United Way of Tarrant County. The selection criteria for inclusion in the sampling frame were identical to Hoefer's (2008). Although not all agencies were located within the geographic boundaries of Fort Worth, all were identified as agencies serving residents of Fort Worth and/or its immediately surrounding communities. In all, 83 agencies were included in the final list of agencies from the greater Fort Worth area. Surveys were sent to all 83 agencies with a link to the internet-based survey software program Survey

9

Monkey, using SSL technology to ensure confidentiality (see Appendix B). Email addresses were obtained through agency websites. Initially, an email with instructions and the survey link was sent to each agency leaders. After four weeks, those agencies that did not respond received a follow-up email, and four weeks later a third email was sent in conjunction with a phone call. All three emails requested a read receipt, and those agencies that did not return a read receipt received a hardcopy version of the survey with instructions for completion and a pre-stamped envelope. The response rate was 62.7%, with an n=52. The survey was based on a design created by Hoefer and utilized in two separate research studies (2000, 2008). The survey includes a combination of open-ended questions, dichotomous questions, nominal questions, and Likert response scaling divided into three categories: individual/organizational information, agency use of program evaluation, and attitude towards evaluation (see Appendix B). Survey data completed electronically was collected through SurveyMonkey. Data from participants completing hardcopy surveys was manually entered into the program.

Both electronic and hardcopy surveys also included an option to participate in an in-depth interview regarding the decision-making process relating to evaluation and accountability within the agency. Ten interview questions, developed by Richard Hoefer based on his prior research with Dallas agencies (2000, 2008), addressed the decision-making process relating to evaluation and accountability within an agency setting. Of all 52 participants completing surveys in Fort Worth, 40 indicated a willingness to be interviewed. Following survey completion, emails were sent to all 40 agency leaders, with 13 actual interviews scheduled and completed from the Fort Worth area. Requests for interviews were also sent at this time to Dallas agencies that had previously consented to be interviewed. Three interviews were scheduled and completed from the Dallas sampling frame. All interviews were conducted by phone using the semi-structured interview format found in Appendix C. The researcher collected interview data by typing key-word notes for each question, reflecting these notes back to the interviewee, modifying and making additions to the response as necessary, and then proceeding to the next question. After

10

asking all ten questions, interviewees were asked if they wished to contribute any additional information to the interview, and following these final statements participants were given an opportunity to debrief with the researcher regarding the research purposes and use of the data.

Survey data was analyzed using the PASW Statistics 17 for Windows software. Survey data from the greater Dallas area was obtained in an SPSS data file from Hoefer. Survey data from the greater Fort Worth area was obtained as an Excel spreadsheet from SurveyMonkey and transferred into an SPSS data file in combination with previously obtained data.

A grounded theory approach was used to analyze interview content using an outline based on Glaser's style of emergent methodology (Dick, 2005). Grounded Theory is a methodology popularly used in theses and dissertations through systematic collection of data in order to produce multivariate conceptual theories. Grounded Theory is considered an empirical methodology that is modifiable, as many verificational methods may rely on poor data, and all data, regardless of its quality, can constantly modify a theory through comparisons (Glaser, 1999). This method, designed to allow the data to organize the research process without the restrictions of a theory, therefore serves to explore the process of how nonprofit organizations make decisions regarding evaluations, rather than test a specific theory about this process (Lingard, Albert and Levinson, 2008).

First, the interviews were read independently, noting frequently mentioned themes. This provided the foundation for a second reading, in which specific themes were identified and documented. A third reading solidified these themes and served to create a list of themes. This list of themes was then reviewed separately to synthesize duplicate themes and resolve any inconsistencies. A revised list of themes was then organized into categories and sub-categories. At various points through each reading, the researcher engaged in a constant process of memoing, jotting down any apparent relationships between variables in a separate document. The list of memos was compared with the list of categories and sub-categories, and then

11

organized to identify trends illuminating the differences in decision making regarding evaluation among Dallas and Fort Worth nonprofit organizations.

CHAPTER 4

RESULTS

An initial analysis reveals that differences do exist between nonprofit agencies in the

Dallas and Fort Worth areas. Caution is extended to the interpretation of these results due to

the differences in sample size and response rate between the greater Dallas area (n=33) and

the greater Fort Worth area (n=53).

4.1 Characteristics of Respondents as Individuals

As shown in Table 4.1, the information seems to come from a higher proportion of

executive directors in Fort Worth, with 80.8% of respondents who were the executive director or

president of the organization; and 7.7% were other top administrators. In comparison, only 69%

of Dallas respondents were the executive director or similar position, with 24.2% being another

type of top administrator. None of these differences are statistically significant.

Table 4.1 Characteristics of Respondents as Individuals

	Dallas (n=33)	Fort Worth (n=52)
Respondents who were executive director/president of the organization	69%	80.8%
Respondents' mean length of service at the organization	7.9 years	12.7 years
Respondents' mean length of service at present job	7.5 years	10.4 years
Respondents' mean age	52.0 years	54.5 years
Percent female (male)	75% (25%)	53.8% (46.2%)
Percent race: Anglo	84%	87%
Percent race: African-American	9%	6%
Percent: Hispanic	3%	4%
Percent: Asian	3%	2%
Percent Race: Other	0%	2%

The mean length of service at the organization of the respondents in Dallas was 7.9

years, with an almost equivalent mean length of service in their present job at 7.5 years. In Fort

Worth, on the other hand, respondents reported their mean length of service at the organization

13

as 12.7 years, with a mean length of service in their present job at a slightly lower 10.4 years. None of the differences between lengths of services reported were statistically significant.

The mean age of respondents in both areas was very similar, with Dallas respondents presenting a mean age of 52.0 years and Fort Worth respondents at 54.5. In Dallas, 75% of respondents were female (25% male); in Fort Worth, only 53.8% of respondents were female (46.2% male). (This is not a statistically significant difference).

In Dallas, 84% of respondents were Anglo, 9% were African-American, 3% were Hispanic and 3% were Asian. None were "Other." In Fort Worth, 87% of the respondents were Anglo, 6% of respondents were African-American, 4% were Hispanic, 2% were Asian and 2% were "Other." There was no statistically significant difference in the distribution between the two cities in question.

4.2 Characteristics of Respondents as Organizations

As shown in Table 4.2, the median agency budget in Dallas was $1950, while in Fort Worth it was $1200. While no statistically significant difference was found to exist between the means of these two groups, further testing revealed that the distributions in the two groups were statistically significantly despite the relatively small difference between medians (Mann-Whitney $U = 516$, p<.05).

Table 4.2 Characteristics of Respondents as Organizations

	Dallas (n=33)	Fort Worth (n=52)
Median Agency Budget (in 000s)	$1950	$1200
Organizational Budget Percent: Government	40%	27%
Organizational Budget Percent: United Way	8%	4%
Organizational Budget Percent: Foundation	17%	13%
Organizational Budget Percent: Other	36%	55%
Mean Organizational Age	41.9 years	44.6 years
Is organization "faith-based"?*	16%	50%

Mean organizational budget percents in Dallas were 40% from government sources, 8% from United Way, 17% from foundations, and 36% from other. These percentages in Fort Worth were 27% from government, 4% from United Way, 13% from foundations, and 55% from other. While the category "Other" consisted of various combinations of funding sources, 43% included churches. Although the data indicate that government sources are more important in Dallas and other funding sources are more frequent in Fort Worth, the differences are not statistically significant.

The average age of the organizations responding in Dallas is 41.9 years, with a range between 8 years and 120 years. The average age of organizations in Fort Worth is 44.6 years, with a range between 6 years and 165 years. These differences are not statistically significant.

Organizations were asked if they considered themselves "faith-based." A comparison between both agencies revealed a statistically significant difference between the proportion of faith-based agencies in each city, with 16% of organizations in Dallas identifying as faith-based in contrast to 50% of organizations in Fort Worth (X^2=10.05; p=0.002).

4.3 Prevalence and Aspect of Evaluation and its Use

Participants answered questions about their use of evaluation, the rigor of the evaluation designs utilized, and the types of measures employed in the evaluation. Significant differences existed in the number of programs conducting evaluation; the influence of government funding on reasons for evaluating, information released to stakeholders; use of evaluations; and the influence of United Way funding on conducting evaluations. Overall, however, organizations in Dallas and Fort Worth are quite similar in the various aspects of evaluation and its use.

4.3.1. Evaluations Conducted

A larger percentage of agencies in Dallas are conducting evaluations in comparison to Fort Worth agencies. In Dallas, 84.8% (n=28) of respondents reported that their largest program has been evaluated in the last two years, a statistically significant difference in comparison to

15

only 57.7% (n=30) of respondents in Fort Worth (X^2=5.832, p=0.016).

Interestingly, a second way of measuring this concept revealed very different results. Respondents were provided with the statement, "Evaluations are conducted regularly in this organization," and asked to disagree or agree. On a scale of 1 to 7, with 1 representing "Very much disagree" and 7 representing "Very much agree," no statistically significant difference exists between mean responses from Dallas agencies (6.39) and Fort Worth agencies (6.18).

Respondents that did not report conducting a program evaluation in the last two years were asked to identify the reasons why an evaluation had not been conducted. The top reason reported by Fort Worth agencies was "A lack of funds to pay for an evaluation" (65%, n=13), whereas Dallas agencies reported "There is no need for an evaluation" (100%, n=5). These numbers are too small to conduct any separate analyses on.

4.3.2   Reasons for Evaluation

The reasons why agencies are conducting evaluations are fairly consistent between organizations in Dallas and Fort Worth. In both cities, the two most common reasons for conducting evaluations are to know how a program is doing and to show funders/potential funders how successful the program is. In Dallas, the most frequent reason for conducting evaluations was tied between curiosity reasons, the desire to know how the program is doing, and commercial reasons, to show funders and potential funders how successful the program is (81%, n=26). Fort Worth respondents cited curiosity as the most common reason to conduct an evaluation (96%, n=28), with commercial reasons coming in second at 79%. Coercion, being required to conduct an evaluation by a funder, was a reason for evaluation for 77% of Dallas agencies but only 68% of Fort Worth agencies. Control, ensuring that the program complied with program procedures, was a reason for evaluating in only 54% of Dallas agencies compared to 75% of Fort Worth agencies. Other reasons were cited by 12% of agencies in Dallas and 25% of agencies in Fort Worth. None of the differences between reasons for conducting evaluations are statistically significant.

4.3.3 Activities Involved In Evaluation

Of the organizations that conducted an evaluation in the prior two years, implementation monitoring was conducted by 74% of Dallas agencies and 68% of Fort Worth agencies. Outcome evaluations were overwhelmingly completed by both agencies in Dallas at 93% and agencies in Fort Worth at 96%. Process evaluations were completed in 85% of Dallas organizations as compared to 61% of Fort Worth organizations. Other evaluation strategies were employed by 15% of Dallas agencies and 4% of Fort Worth agencies.

4.3.4 Sources of Expertise

The sources of evaluation expertise are similar for both Dallas and Fort Worth organizations. While both cities rely predominantly on a combination of internal and external sources of expertise, the dominance of internal-only to external-only sources of expertise indicates a possible trend towards institutionalized evaluation in both cities.

In Dallas agencies, the sources of expertise for the evaluation were: 33% inside the organization, 15% external to the organization and 52% from both inside and outside the organization. In Fort Worth, the sources were: 29% inside the organization, 4% external to the agency and 68% from both inside and outside the organization. It appears that many agencies in Dallas are more likely to rely on external sources of evaluation expertise than agencies in Fort Worth, while organizations in Fort Worth are more likely to rely on a combination of internal and external expertise than organizations in Dallas.

In both Dallas and Fort Worth, funding is most likely to come from internal funding sources (63% in Dallas and 75% in Fort Worth). The second most frequent source of funding is program grants, with almost twice as many Dallas organizations (44%) relying on these as a funding source in comparison to 25% of Fort Worth organizations. 11% of respondents in Dallas indicate the use of separate grants in contrast to only 4% of Fort Worth respondents. Donated labor is used as a funding source for evaluations in 11% of Dallas organizations and 7% of Fort Worth organizations. Only 4% of Dallas agencies report other funding streams, while Fort Worth

17

agencies report no other funding sources for conducting evaluations. Overall, it appears that Dallas and Fort Worth respondents are fairly similar in the frequencies with which they rely on different types of funding streams to conduct evaluations. (Note that more than one type of funding stream could be checked, so totals add up to more than 100%).

4.3.5 Evaluation Design

The types of evaluation designs utilized are similarly rigorous in both Dallas and Fort Worth, with the pretest-posttest design as the most frequently used method, followed by a time series design and a posttest only design. Experimental evaluations, the most rigorous of all designs, compose a relatively small percentage of all evaluations conducted in both Dallas and Fort Worth.

In Dallas, posttest only was used by 15%; pretest-posttest by 59%; time series by 33%; comparison posttest by 4%; comparison group pretest-posttest by 0%; and any type of experimental design by 7%. In Fort Worth, posttest only was used by 22%; pretest-posttest by 41%; time series by 41%; a comparison posttest by 0%; a comparison group pretest-posttest by 11% and any type of experimental design by 7%. (Note that the wording of this question was modified to allow participants to make finer distinctions in the more rigorous evaluation designs).

4.3.6 Source of Evaluation Data

Data sources used in evaluations appear to be equally rigorous across organizations in both the Dallas and Fort Worth areas. In both cities, the major source of data is agency records, followed by non-standardized instruments. Standardized instruments, among the most rigorous data collection methods, also appear to be a significant data source in evaluations.

In Dallas, agency records were utilized in 85% of agencies; non-standardized instruments were used by 63%; standardized instruments by 52%; and social indicators by 19%. In Fort Worth, 78% of agencies reported the use of agency records; 63% the use of non-standardized instruments; 37% the use of non-standardized instruments; and 19% the use of social indicators. In both Dallas and Fort Worth, over three quarters of all agencies rely on

agency records, and over half rely on non-standardized instruments. Standardized instruments are utilized by over half of agencies in Dallas and one third of agencies in Fort Worth.

4.3.7 *Use of Evaluation Results*

Respondents in Dallas and Fort Worth are using the results of program evaluation for various purposes with similar frequency, with the top three purposes as program improvement, goal achievement, and advocating for new funds.

In Dallas, 100% of all agencies report using the results of program evaluation for program improvement; 89% for verifying goal achievement; 7% to shift funding away from the program being evaluated; 22% to shift funding to the program being evaluated; 78% to advocate for more money from original funders; and 78% to advocate for more money from new funders. Fort Worth organizations report that 100% use program evaluation for program improvement; 96% use it for goal achievement; 11% to shift funding away from the program; 19% to shift funding to the program; 70% to advocate for more money from the original funders, and 63% to advocate for more money from new funders.

4.3.8 *Final Arbitrator*

In both Dallas and Fort Worth, nearly three fourths of all agencies report that the executive director is the final arbitrator of how much of the evaluation's conclusions and recommendations were released and to whom. Only one agency reported the evaluator as the person making the final decision.

Among Dallas agencies, executive directors make the final decisions for 74% of agencies, similarly to 78% for Fort Worth agencies. In the remainder of Dallas agencies (26%), the organization's board of directors made the final decision, while no Dallas organizations reported the evaluator making the final decision. Among Fort Worth agencies, the board of directors made the final decision 15% of the time, while 4% of organizations actually reported the evaluator holding the final decision.

19

4.3.9 Dissemination of Information

Following a decision regarding which information to release, respondents were asked to whom the information was shared. Organizations in both cities share information on very similar levels, with managers and the agency's board of directors as the top two recipients of information.

In Dallas, 85% of respondents formally shared the information with managers, 85% with front-line staff, 96% with the agency's board of directors, two-fifths (39%) with the public via publication on a website or other form, 85% with funders, and 19% with press or other media. For Fort Worth agencies, 100% reported sharing the information with managers, 85% with front line staff, 89% with the agency's board of directors, two-fifths (41%) with the public via publication on a website or other form, 85% with funders, and 15% with press or other media. One agency in Fort Worth also reported sharing information to the public through the website of an outside regulatory agency.

<u>4.4 Influence of Funding Streams</u>

4.4.1 Influence of Government Funding

In order to assess whether funding sources impact the use of evaluation and accountability strategies, Dallas and Fort Worth organizations were examined to determine whether the receipt of government funding made a difference. Results indicate that the receipt of government funding may have a differing influence on the Dallas and Fort Worth areas. Among Fort Worth agencies, a higher receipt of government funding indicates a higher degree of coercion as a reason for evaluating as well as a decrease in the release of evaluation information to the public. Among Dallas agencies, higher government funding appears to be related to an increased use of evaluation to advocate for money from the agency's original funders.

The median percentage of government funds in organizational budgets for both groups of respondents is 30%. Dallas and Fort Worth were separately evaluated to determine whether

the receipt of government funding made a difference. No significant differences were found in either Fort Worth or Dallas regarding the effect of government funding in the following areas: sources of expertise for an evaluation, funding sources for an evaluation, types of activities encompassed in an evaluation, evaluation design, type of data used in an evaluation, and person making final decision regarding evaluation data.

In considering reasons for evaluating, 84% of Fort Worth agencies with higher government funding listed "coercion," in comparison to 16% of Fort Worth agencies with less government funding. This difference was statistically significant ($X^{2=}13.669$, p=0.000), while Dallas had no significant differences in this area.

84% of Dallas agencies with higher government funding listed "advocating for more money from original funders" as a use for evaluation, in comparison to 16% of Dallas agencies with less government funding. This difference was statistically significant ($X^{2=}5.051$, p=0.025), while Fort Worth had no significant differences in this area.

Regarding which stakeholders receive information from the evaluation, 73% of Fort Worth agencies with less government funding indicated releasing information to the public, in comparison with only 17% of those with higher government funding. This difference was statistically significant ($X^{2}=7.116$, p=0.008), while no significant differences were found in Dallas. A regression analysis conducted on Fort Worth organizations confirmed that an increase in government funding was associated with a decreased frequency of conducting evaluations (B=.011, standard Beta=.385, p=.047), a greater likelihood to report coercion as a reason for evaluating (B=.011, standard Beta=.755, p=.00), and a decreased release of information to the public (B=-.008, standard Beta=-.522, p=.006). These relationships were not found to exist in Dallas at a significant level.

4.4.2 Influence of United Way Funding

The receipt of United Way funds was found to be a predictor of the use of program evaluation. The location of the survey, however, was not. The distribution of agencies receiving

United Way funding is fairly consistent in both Dallas and Fort Worth, with 51.6% agencies in Dallas received United Way funding in comparison to 46.9% of agencies in Fort Worth. Holding the location of the survey constant, the receipt of United Way funds was associated with a statistically significant increase in the use of program evaluation across organizations in both cities (B=0.668, standardized Beta of .321, p=.018). However, when looking at Dallas and Fort Worth separately, the receipt of United Way funds was associated with a statistically significant increase in the use of program evaluation in Fort Worth, but not in Dallas (B=.710, standardized Beta of .392, p=.043). The location of the survey, when controlling for the receipt of United Way funds, did not have a significant impact on the use of program evaluation. (Receipt of United Way funds was dichotomous because United Way requirements for outcome evaluation apply either totally if funding is received or not at all.)

4.5 Influence of Faith-Based Status

Given the significant differences in the number of organizations professing to be faith-based, organizations in Dallas and Fort Worth were also examined to determine the impact of identification as faith-based on conducting evaluations. In Dallas, no significant relationship existed between faith-based status and conducting evaluations. In Fort Worth, however, organizations that identified as faith-based were associated with a statistically significant decrease in the use of program evaluation (B=-.744, standardized Beta of -.401, p=.034).

Overall, it appears that more Dallas organizations are conducting evaluations in comparison to Fort Worth agencies. For the most part, Dallas and Fort Worth are fairly similar in their various aspects of evaluation, with some differences in the type of influence exerted by governmental funding streams and United Way funding.

4.6 Attitudes about Evaluation

In keeping with the similarities between Dallas and Fort Worth agencies as noted above, attitudes about evaluation are fairly consistent. Thirteen statements were presented to respondents who were asked to rate how true they were on a seven point scale, with a low

score (1) indicating that they very much disagreed and a high score (7) indicating that they very much agreed. Four of these statements can be seen as positively regarding evaluation, while nine can be said to be negative (shown in Table 4.3). Responses to each question were averaged, and t-tests were conducted to compare mean responses in both cities. Differences between Dallas and Fort Worth were found to be small and statistically insignificant.

Table 4.3: Mean Attitudes about Evaluation* in Dallas and Fort Worth

PART A: Positive Statements about Evaluation**	Dallas (n=33)	Fort Worth (n=52)
Broadly speaking, evaluations are useful.	6.29	6.25
An evaluation can help improve the program.	6.57	6.21
An evaluation can inform funders of good things the program does.	6.46	6.50
The cost of an evaluation is money well-spent.	5.36	5.67
Evaluations are conducted regularly within this organization.	6.39	6.18
PART B: Negative Statements about Evaluation**		
Evaluations are usually biased or inaccurate.	2.32	2.00
The strengths and weaknesses of the program are already known.	3.50	3.43
Evaluations can hurt funding if the results are negative.	3.39	3.54
An evaluation is too much work.	2.96	2.36
Evaluation results are generally of no use to direct practice staff	1.68	1.79
Evaluations merely state the obvious.	2.14	2.50
Evaluations are inherently more trouble than they are worth.	2.04	2.07
No evaluation is needed because we already know that the program is doing a good job.	1.82	2.04
If the results of an evaluation are negative, the evaluation is probably flawed.	1.86	2.32

Possible responses are: 1 (Very much disagree); 2 (Fairly much disagree); 3 (Slightly disagree); 4 (Neither disagree nor agree); 5 (slightly agree); 6 (Fairly much agree); and 7 (very much agree).
*** None of the differences in Table 4.6 were statistically significant.*

4.7 Decision-Making in Areas of Evaluation and Accountability

Interviews were conducted to explore the decision-making process regarding evaluation within agencies. Although the survey asked several questions regarding the decision-making process surrounding evaluations (e.g. "Please mark all the reasons why an evaluation has been conducted"), Hoefer's prior research made it clear that much remained to be known regarding

the decision-making process relating to evaluations and accountability within an agency (1999, 2008).

In order to examine this decision-making process, participants were asked a series of open-ended questions exploring the following areas:

- Concepts of accountability

- Aspects of evaluation process

- Dissemination of evaluation information

- Conceptual and practical applications of evaluations in regards to accountability

- Influence of funders in conducting evaluations

- Perceptions of future of evaluations

A significant discrepancy existed in sample sizes of respondents (Dallas n=3, Fort Worth n=13), and no significant differences were found between both groups of respondents. Due to these factors, results are presented as an aggregate of participants in both locations.

4.7.1 Concepts of Accountability

Participants were asked to define their individual and organizational concepts of general accountability, accountability to society, and accountability to clients. Various themes overlapped in these three areas as participants described what accountability meant to them. All participants reported between three and six properties in their concepts of accountability. The most prevalent property of accountability included fiduciary responsibility, reported by 81% of all organizations (three in Dallas and ten in Fort Worth). The second most prevalent property was integrity, referred to most often as "doing what we said we are going to do," with 75% of all agencies citing this as a property of accountability (two in Dallas, ten in Fort Worth). The third most frequent category was quality, encompassing quality of service, delivery, and/or outcome, with 69% of organizations mentioning this (two in Dallas, nine in Fort Worth). Interestingly, only

24

25% of all agencies reported the act of conducting evaluations as a property of accountability (three Dallas agencies, three in Fort Worth). Other properties of accountability include:

- o Agency mission as contract with individuals and society

- o Benevolent/charitable nature of work

- o Critical self-evaluation of effectiveness

- o Client ownership within agency programming

- o Staff competency

- o Leadership among other organizations

- o Guidance of a higher power

In asking agencies to describe their concept of accountability to society, differences arose in participants' definitions of "society at large." 38% of all agencies (two in Dallas, one in Forth) envisioned society at large as actual and potential stakeholders in their programs. 31% of all agencies (five Fort Worth agencies) described society at large as actual or potential clients, and 25% (one Dallas agency, three Fort Worth agencies) viewed society as large as all individuals within a specific geographic area of service.

4.7.2 Decisions in Evaluation Process

Participants were asked to describe the process of conducting evaluations within their organizations, from the design to the implementation of the evaluation. Their responses were then analyzed in order to identify trends in decision-making during the actual evaluation process. From this process, two distinct trends emerged along different stages of the process.

In the design process, agencies differed distinctly in the individuals involved in the design of the evaluation. Evaluation planning that involves front-line staff and/or clients incorporates several of the elements of accountability as described by participants (client ownership, critical self-evaluation, staff competency). Looking at all the agencies combined, only 31% (n=5), one from Dallas and four from Fort Worth reported the involvement of front-line staff and/or consumers in their evaluation design. The remaining 69%, two from Dallas and nine

from Fort Worth, involved various combinations of top-line staff and evaluators. Some agencies describe their reasoning for such decisions:

> "We involve clients, focus groups, and work with evaluators to break down services to a level that can be measured so we can actually see we are on target…We talk [with staff] about strategies, case load size, so it's very fluid at that point about making decisions on how and what's going to be evaluated." (Participant 16)

The second trend identified existed in organizational decisions of individuals involved in collecting evaluation data. The great majority, 69% (n=11), two from Dallas and nine from Fort Worth, relied on front-line staff to collect data, and 18% (one from Dallas , two from Fort Worth) relied on non-program staff. In one Fort Worth agency, an internal evaluator directly collected data. Due to a unique relationship with a legal entity, another Fort Worth agency relied on an external evaluator to collect data for its program evaluation. Some participants utilizing front-line staff describe the reasoning for their decision:

> "[Follow-up] phone calls are typically made by people that [the clients] used to know, so even if we want to be sensitive to bias, folks are usually able to share how positively or badly they felt as a result of the intervention." (Participant 12)

Other agencies, however, affirm that their decision to use non-program staff is the best option for data collection:

> "Surveys are distributed and collected by the receptionist….We want to be transparent so it is not part of the clinician's practice or exchange with information. Surveys weren't getting done by therapists because they have so much responsibility and they were losing too much data." (Participant 3)

Additionally, one participant commented on the unique role of an internal evaluator as a data collector:

"By working closely with the client from the beginning, [the evaluator] had built a relationship that had engaged so that at follow-up he was able to find those people. Even if clients became homeless, they liked him and they kept him informed whenever they moved around." (Participant 16)

4.7.3 Decision Making in Dissemination of Information

Hoefer (2000) states that one measure of organizational accountability is whether the results of an evaluation are being released to front-line staff. Survey results indicate that, in both Dallas and Fort Worth, 85% of survey respondents release information to front-line staff. Data collected from interview respondents, however, identifies that the degree to which this occurs varies dramatically between agencies. 69% of all respondents (two in Dallas, nine in Fort Worth) report that data is formally shared with front-line staff through weekly meetings, formal reports, and/or informal conversations, while 31% specify no mechanism through which evaluation results are made available to front-line staff (one in Dallas, four in Fort Worth).

Of the eleven agencies that release information to staff, 36% (two in Dallas and one in Fort Worth) involve front-line staff in the improvement of program performance and 28% (one in Dallas, two in Fort Worth) produce employee-specific data to guide a worker's individual efforts. However, the remaining 36% (one in Dallas, four in Fort Worth) do not incorporate staff in improving program or individual performance.

"Staff use reports to adjust focus of their energy…and they many need to reshift their energy in a different direction." (Participant 16)

"Identifying overall [performance] in aggregate compared to an individual person helps identify a [programmatic] problem and creates a training issue." (Participant 13)

A total of ten participants reported that evaluation results were made available to the community (two from Dallas, eight from Fort Worth), although the means described for such distribution varied drastically. The tools reported as carrying out this dissemination of

27

information included: agency websites, annual reports, emails, tours, presentations, special events, newsletters, social networking websites, and published articles. Considering the previously mentioned discrepancies in concepts of "society at large," further data is needed to explore which communities are being targeted for information (clients, providers, funders) , and how organizations tailor their means of disseminating evaluation data based on the needs of community stakeholders.

4.7.4 Use of Evaluation in Establishing Accountability

Participants were questioned about their views on evaluation as a means of establishing accountability in order to identify their perceptions about the relationship between evaluation and accountability. Four major themes emerged from participant responses: integrity, fiduciary accountability, quality, and mission. All four themes coincide with properties of accountability identified previously; in fact, the first three themes listed reflect the three most frequent properties of accountability described by participants.

4.7.4.1 Evaluation as Integrity

The most frequent reason cited by 56% of all organizations (two from Dallas, seven from Fort Worth) as the manner in which evaluations establish accountability is the creation of integrity. This concept is is frequently described as "proof that we did what we said we were going to do." Aspects of evaluation establishing accountability through integrity include recognizing the limitations of anecdotal data; the ability to create measurable milestones; measuring outcomes rather than outputs; and the proper implementation of a program model.

4.7.4.2 Evaluation as Fiduciary Accountability

One quarter of participants (25%--all four in Fort Worth) reported that evaluations established accountability through fiduciary disclosure. Specifically, evaluations were described as assisting in decisions to make the best use of dollars by shifting funding to/away from programs and providing evidence of the proper use of funds.

4.7.4.3 Evaluation as Quality

19% of participants (three in Fort Worth) reported that evaluations helped enhance the quality of services and outcomes, therefore improving accountability. Specifically, evaluations were described as a factor in staff discussions about improving services, assessing the degree to which the services impacted the identified need, and expanding programs to better fit with reported client needs.

4.7.4.4 Evaluation as Mission

13% of participants (one from Dallas, one from Fort Worth) described evaluation as a tool in achieving accountability through evaluation's contribution to fulfillment of organizational missions and visions. Specifically, evaluation was viewed as creating measurable milestones towards the agency's vision, creating community support for the agency mission, and creating a "culture of evaluation" that permeated all activities.

4.7.4.5 Examples

As a further means of evaluating whether agencies were actually utilizing evaluations in decisions to increase organizational accountability, participants were asked to provide a specific example illustrating this fact. One organization (Dallas) was unable to articulate any examples. Overwhelmingly, 63% of all organizations (two in Dallas and eight in Fort Worth) reported improvements in program services as a result of evaluations. Of the remaining four organizations (all in Fort Worth), 13% reported using evaluation data in making financial decisions, 13% reported using evaluation data to enhance and improve evaluation design, and 6% reported using evaluation data to provide leadership by serving as a consultant to another organization.

4.7.5 Role of Coercion

Participants were asked to describe how coercion from funders influenced their organization's decision to conduct evaluations. The results indicate that organizations encounter coercion from funders in both their decisions to evaluate and the evaluation design.

4.7.5.1 Coercion's Role in Decision to Evaluate

38% of organizations, all from Fort Worth, reported coercion from funders as having a significant role in their decision to evaluate. All participants were varied in their general attitudes toward coercion, with some viewing coercion negatively while others viewed it in a more positive light.

> "Ugh…Grant makers are wanting so much more…The majority of the time the decisions [to evaluate] come because of the grants…[One grant] wants us to capture information that's] cumbersome and terrible, then you have to configure that to the program design." (Participant 11)

> "I doubt we would have done program evaluations if [one of our funders] hadn't said you have to do them…You want to prove to funders that you are doing a good job, and this is a way of doing that…I wish more funders would coerce us, but I think that's a way of proving you are." (Participant 15)

4.7.5.2 Coercion as Factor in Evaluation Design

63% of all participants, three from Dallas and seven from Fort Worth reported that while coercion did not play a significant role in their decision to evaluate, coercion from funders frequently arose in the design of the evaluation. Particularly, all participants reported experiencing coercion from funders in the types of measures they were being asked to evaluate. Organizations reported dealing with this type of coercion through two main strategies: acceptance and negotiation.

Of the five agencies (31%) reporting acceptance of coercion in their evaluation design, one was from Dallas and four from Fort Worth.

> "With a grant you have to give [the funder] goals and outcomes, which might be different than something they individually had. You may end up measuring something in addition. It's just what you do." (Participant 4)

"The coercion is that they ask us to measure things that we don't consider important, such as a focus on outputs rather than outcomes to satisfy other stakeholders…with very little attention to whether…it impacts the problem it's designed to impact." (Participant 2)

Of the remaining six agencies, two in Dallas and three in Fort Worth, organizations appear to rely on two main heuristic methods in their decisions to modify their evaluation design: fit with mission/values and integrity of program:

"If an agency is getting coerced around a certain evaluation strategy, they need to evaluate why they are doing that…It may be a judgment call that what you're measuring doesn't fit with the agency's mission…and pulls away from your ability to provide quality services." (Participant 16)

"It's a waste of time to evaluate programs in ways that are not meaningful. If you're doing it only because a funder asks it's worthless." (Participant 13)

4.7.6 Evaluation as Giving an Edge in Seeking Funds

Participants were asked whether they felt that conducting evaluations gave them an "edge" in conducting evaluations. None of the participants reported that their evaluation compared unfavorably to their competitors, describing themselves in terms such as "favorable," "on top of the game," and "ahead than most." All participants agreed that conducting evaluations gave them an edge, although the degree to which this was expressed varied significantly. Most interestingly, however, were the reasons described for how evaluation gave them an edge. These reasons can be summarized in rational and political terms.

Rational reasons can be described as the organization conducting the evaluation to comply with funders in rational goals, such as pursuing accountability through evaluations. Six participants (one from Dallas, one in Fort Worth) described getting an "edge" from evaluations because they demonstrated the rational goals of program effectiveness through empirical data, and demonstration that funds were well-spent.

31

"[Through program evaluations], we can speak empirically [to funders] about
the results we have achieved, the ones we are going to achieve, and how we
are going to assess success or failure of the program." (Participant 1)

Political reasons are understood to be aspects of evaluation conducted for various pragmatic
reasons outside of a rational means of establishing accountability. Ten participants (two in
Dallas, eight in Fort Worth) reported that their "edge" from evaluations was a result of attaining
the political ends of competition and social status.

"In today's environment, if you're not trying to accomplish, you're not a high-
level performing nonprofit…[If you don't], just like lots of businesses that are
mediocre, in bad times, they go out of business." (Participant 16)

4.7.7 Trends in Evaluation

Participants were asked their opinions regarding the future of evaluations. Their
responses were analyzed and broken down into three categories: factors influencing the future
of evaluations, anticipated changes in evaluation design, and barriers in responding to these
changes. In reviewing these categories, a consistent picture emerged as responses converged
to create a picture of the future of evaluation in Dallas and Fort Worth.

The three main environmental variables perceived as having an impact on the future of
evaluations include competition, collaboration, and knowledge. Competition is seen as a force
that will continue to drive evaluations as a result of decreases in funds/funders, increases in the
number of nonprofits, increases in the number of funders asking for evaluations, and the
growing use of evaluation as a negotiation tool with funders. Collaboration is also seen as a
powerful new trend, as funders increasingly ask agencies to demonstrate partnerships to create
more effective programs. Knowledge is perceived as a powerful factor, as more agencies
become knowledgeable in doing evaluations, evaluations being conducted grow stronger,
evaluations are increasingly seen as a means of establishing accountability, and agencies
develop pride in their ability to conduct evaluations.

Participants indicate anticipation that evaluation designs will become increasingly complex as a result of the variables described above. It is anticipated that funders will increasingly demand evidence-based models and rigorous research methods to evaluate programs; improved program planning through logic models and program descriptions; an increasing focus on outcomes, particularly long-term outcomes combined with a decreasing focus on outputs; and a growing push to link outcomes, rather than outputs, to expenditures.

Participants also identified several barriers to successfully responding to the anticipated changes in evaluation. Primarily noted was the lack of expertise to respond to the increasing rigor of evaluations. Many organizations described barriers within their level of internal expertise, citing a need for increased knowledge in designing evaluation plans, developing measures for prevention programs, and creating measurement tools. Agencies also reported a need for external expertise, describing the poor responsiveness from educators in programs of higher education; the deficiency of such programs in preparing agency leaders for program evaluation; insufficient research regarding evidence-based interventions, and gaps in the application of such research to practice.

Another barrier faced by organizations in preparing for the future of evaluations is the priorities set by funders. Funders are often perceived as setting outcomes that don't yield useful information and failing to allocate funds for evaluation purposes. Additionally, some organizations report that funders are negatively influenced by a political system that tends to overlook stigmatized clients and prefers to fund intervention results in lieu of prevention efforts.

CHAPTER 5

DISCUSSION

Through 85 survey respondents and 16 interviews, nonprofit agencies in Dallas and Fort Worth were compared to yield insight into any significant differences regarding evaluation and accountability. Results indicate that the prevalence of faith-based organizations in Fort Worth is a significant predictor of in conducting program evaluations. Fort Worth is home to significantly more faith-based organizations than Dallas. At the same time, Dallas organizations are conducting evaluations at a significantly higher rate than agencies in Fort Worth. Interview data suggests that churches as funders of faith-based organizations tend to perceive giving as an act of faith, and thus may view program evaluations as questioning the higher power that motivated them to give. More data is needed to explore potential differences between Dallas and Fort Worth in the nature of faith-based organizations and the role of faith within such organizations. Further research is also needed to examine the relationship between conducting evaluations and the percentage of funds obtained from faith communities. Testing should be conducted to evaluate the influence of faith-based status on any other significant concepts of evaluation between Dallas and Fort Worth.

Dallas and Fort Worth organizations were found to be quite similar in their attitudes about program evaluation. The lack of significant differences can certainly be viewed as a positive, as this might be indicative that attitudes about evaluations among Fort Worth organizations have improved over time, much like Hoefer demonstrated for Dallas organizations (2008). This would suggest that the improvement in attitudes about program evaluation is due to variables not restricted to either the Fort Worth or Dallas areas, but potentially reflects regional or broader geographic trends. Lacking historical data for comparison among Fort Worth

34

agencies however, this requires further investigation to yield longitudinal data that can be used in further analysis.

No significant differences existed between organizations Dallas and Fort Worth in actually conducting program evaluations, types of evaluations conducted, use of design and data sources, and purposes of evaluations. However, it is interesting to observe how the influence of government funding differentially affects both cities.

There is not a statistically significant difference between the percent of mean organizational percent of government funding in Dallas and Fort Worth (40% in Dallas, 27% in Fort Worth). Despite this fact, significant differences arose between agencies in Dallas and Fort Worth when comparing those above and below the median percent of government funding. In Fort Worth, higher levels of government funding were significantly associated with an increase in coercion from funders as a reason for evaluating and a decrease in releasing information to the public. Additionally, among Fort Worth agencies an increase in government funding was also found to be associated with a decrease in conducting evaluations. As cited earlier, Frumkin and Kim (2002) theorized that government contributions, typically awarded for programmatic initiatives, lead to decreased administrative efficiency due to large amounts of administrative overhead rates. In Fort Worth, it is possible that agencies that receive an above-median percentage of government funding are focusing on new initiatives and are therefore less prepared, or less invested, in evaluating programs at early stages. Alternately, Fort Worth agencies receiving significant amounts of government funding might be less inclined to include private stakeholders in their definitions of "society as large". This in turn could lead to a reduced sense of accountability to the general public, culminating in a higher sense of coercion from funders to conduct evaluations along with a reduced sharing of evaluation results. Future research should test for these possibilities by exploring the impact of program age and definitions of society at large in relation to uses of evaluation to establish public accountability.

In Dallas, agencies receiving more than the median percent of government funding reported that evaluations would be used to advocate for more funding from existing funders at a significantly higher rate than those below the median. It appears that Dallas agencies with higher percentages of government funding continue to seek public funds without significant influence on their use of evaluation to establish accountability. More data is needed to explain differences in the influence of government funding on Dallas and Fort Worth nonprofits. Dallas and Fort Worth nonprofits should continue to be rigorously compared to explore differences in specific public funding sources and organizational relationships with those funding sources.

Additionally, among Fort Worth agencies the receipt of United Way funding is associated with a higher likelihood to conduct evaluations. Further testing should be done to examine the relationships of government funding and United Way funding to agency size, years in existence, and other demographic variables of respondents that may yield further insight into the observed differences. Future research should continue to explore the influence of faith-based, governmental and United Way funding streams.

Sixteen participants, three in Dallas and thirteen in Fort Worth, participated in interviews that discussed various elements of decision-making relating to evaluation and accountability within their agencies. While the sample size was too small to make many strong comparisons, an examination of the participants as a whole revealed various aspects of decision-making. Agencies varied in their definitions of accountability, with only 25% of agencies mentioning evaluation as one of these properties. This suggests that the manner in which an agency defines accountability and to whom it owes that accountability may influence its likelihood to conduct evaluations. More testing is needed to examine the relationships between concepts of accountability, society at large, and the frequency of conducting evaluations.

Two trends observed as part of the evaluation process include the people involved in evaluation design and data collection. Despite differing views on the role of clients and front-line

36

staff in the evaluation process, the involvement of either of these parties in the evaluation design has the potential to increase accountability to both clients and society at large through improved responses to programmatic, staffing, and client-related needs. While Hoefer cites the use of rigorous methods as a cornerstone of evaluations in establishing accountability, identified barriers such as the availability of evaluation funds and expertise appear to limit the number of agencies that are actually conducting rigorous evaluations (7%). Participants describe an increasing trend towards both competition and collaboration in seeking funds, indicating that formally engaging clients and staff at deeper levels of the evaluation process may increase an agency's funding potential, therefore enabling it to pursue more rigorous research methods. Future research should specifically examine the frequency of involving front-line staff and/or consumers in the evaluation process, and the corresponding influence on accountability. Considering the prevalence of agency records as a source of evaluation data, careful evaluation planning should strive to balance the empowerment of staff and consumers with effective and efficient data collection methods.

Historically, Hoefer documents that coercion from funders has been perceived as a primary compeller in agency decisions to conduct evaluations (2000). The influence of coercion appears to be changing however, as an increasing number of organizations report the existence of coercion in evaluation design, rather than in their decision to evaluate. Reports are mixed regarding this occurrence: some organizations dislike being asked to measure outcomes considered insignificant, while other organizations do not appear to mind complying with a funder's request if it doesn't relate to the originally stated program purpose. The influence of evaluations in seeking funds is also varied, as organizations report a mix of rational and political incentives for conducting evaluations. Due to the varying perceptions of the role of coercion, coercion as a concept should be further defined and operationalized for future research. Further testing should also be done to test for relationships between perceptions of coercion and the level of expertise of both the funding source and the nonprofit organization. Considering

37

participant predictions that evaluation designs will only continue to grow in complexity, future research should continue to explore the relationships between coercion at all levels of the evaluation process and its influence on accountability.

Results indicate that funder priorities are a significant influence in the future of evaluations. Therefore, an important area for future research is to examine program evaluation and accountability according to types of services. Nonprofit agencies providing services such as sex education and needle exchange programs may rank very differently on a funder's priorities than domestic violence and child welfare services.

Organizations with prevention programs appear to face a unique dilemma in the future of evaluations. With increasing demands for rigorous evaluations, decreasing funds, and a lack of expertise, prevention programs must find secure sources of funding to continue to flourish. One alternative might be to collaborate with local funding entities, such as United Way, to provide technical assistance in developing measurable and meaningful outcomes. Alternately, if funders continue to encourage collaborations, prevention programs would be strong candidates if they organize and negotiate with funders to ensure the availability of evaluation funds.

5.1 Strengths and Limitations

This study faces several limitations, as various threats to both internal and external validity exist. The survey results for Dallas and Fort Worth were collected at two different intervals approximately one year apart, with the small but existent possibility that history and maturation could have affected the results. The differential selection of participants through the volunteer nature of respondents' participation in both the survey and the interview also creates a potential bias. The influence of experimental mortality could affect the validity of the results, as the use of online survey technology with Fort Worth respondents facilitated partial responses through the option of premature survey termination. Additionally, the use of two measures creates a testing threat, as interview responses could have been influenced by previous

participation in the survey. The results of the regression analyses utilized in this research should also be taken with caution considering the small sample size involved in such testing.

Challenges to external validity exist due to the selection of subjects. Selection-interaction threats are present due to the nonrandom selection of participants, as both sampling groups were located in differently-sized cities in an urban area of North Texas and results cannot be generalized to nonprofit agencies in other urban or rural areas. The qualitative data collected through the interviews is also vulnerable to experimenter effects and evaluator bias, while the Hawthorne effect may have influenced respondents' expressed attitudes about evaluation. This one-time collection of also data does not take into account various organizational or contextual variables such as financial difficulties, pending investigations, community support, and environmental circumstances.

Despite these limitations, this is the first known study to empirically compare two networks of nonprofit agencies in their use of program evaluation to establish accountability. Additionally, this study is insightful in examining potential factors influencing the use of evaluation to create accountability while illuminating directions for future research. This study is innovative in combining qualitative and quantitative data to examine the use of evaluation within nonprofit agencies at a more profound level.

5.2 Conclusion

Through a comparison of two urban areas in the North Texas region, this study takes a first step at examining differences in the use of program evaluation to establish accountability. Although this study does not propose to examine the causes for such differences, the results yield useful insights for examining organizational, financial, and conceptual elements that influence the use of program evaluations to demonstrate accountability. A review of Hoefer's four-point criteria to determine if evaluation is being used to establish accountability helps summarize the major findings of this study.

5.2.1 Is Evaluation Being Done?

This study indicates that evaluations are being done in both Dallas and Fort Worth. In order to gain a better understanding of the variables influencing the prevalence of evaluations, future research should focus on exploration of other organizational variables in Dallas and Fort Worth such as: staffing size and programmatic to administrative cost ratios; organizational perceptions of government funders; influence of coercion at all levels of evaluation; and definitions of significant stakeholders.

Of significance are the implications of the relationship between faith-based organizations and evaluations as establishing accountability. Due to the increasing momentum for the growth of faith-based organization in providing social services, social workers must be prepared to collaborate with faith-based organizations by exploring the manner in which these agencies participate in evaluation to establish accountability, and balancing evaluation with the core values of such agencies. Future research should include faith-based organizations through further analysis of relationships between faith-based status and other variables explored in this study, as well as factors such as agency size (including staff and volunteers); relationships with specific religious denominations; expressed religiosity; influence of funding streams, particularly congregational funding; educational backgrounds and expertise of evaluators and key decision-makers; perceptions of evaluation and accountability; and information-sharing with stakeholders.

5.2.2 Are Appropriately Rigorous Research Methods Being Used?

This study demonstrates that organizations in both Dallas and Fort Worth have significant room for improvement in the rigor of their research methods. While various barriers to conducting strong program evaluations are discussed, future research should continue to explore the role of coercion in evaluation design, as well as the influence of funders on the rigor of program evaluations. In light of the existing challenges to utilizing rigorous research methods, organizations have an opportunity to strengthen their accountability through the use of

consumers and front-line staff in the design and implementation of their evaluation process. This practice has the potential to empower clients and staff, create greater awareness of the importance of evaluation, strengthen an organization's competitive edge in seeking funds, and enable an agency to pursue more rigorous research methods in the future.

5.2.3   Are Front-Line Staff Members Learning The Results Of The Evaluation?

The degree to which front-line staff members are learning the results of the evaluation varies between agencies. The establishment of a "culture of evaluation" is critical in allowing agency staff to understand the importance of program evaluation, which in turn encourages high quality data collection. In order to help create this culture, agencies can ensure that evaluation results are not just being made available, but actually discussed and analyzed regularly with front-line staff. Involving staff in this process can improve the likelihood that evaluation results will actually be used to establish organizational accountability Future studies should assess the impact on accountability of involving front-line staff and consumers in the evaluation process, while examining the relationship between accountability and the types of services offered by agencies.

5.2.4. Are Other Stakeholders (Including The Public) Learning The Results Of The Evaluation?

Organizations in both Dallas and Fort Worth are, for the most part, sharing the results of evaluation with the public. As organizations have been shown to differ in their definitions of "society at large," more research is needed to examine who is actually receiving these results, and which means are being used to make those results available. This study also suggests that funding sources may negatively influence the degree to which evaluation results are provided to the public, indicating the need to further examine how funders influence the use of evaluation to establish accountability. In order to ensure that organizations have the necessary tools at their disposal to conduct successful evaluations, future research should recruit organizations from a diverse sample of geographic areas, while pursuing ongoing longitudinal data in Dallas and Fort Worth to analyze ongoing trends and patterns in the use of program evaluation.

41

By involving several social service agencies, this study hopes to generate discussion within existing networks of social service agency providers that can help guide future research into the field of program evaluation and accountability. Through helping agencies improve their evaluation systems to create higher accountability, this study can not only contribute to the sustainability of existing programs, but strengthen the credibility of social service agencies and stimulate greater access to social services.

The recent economic downturn and financial scandals are a keen reminder of the importance of public accountability. One way that social work can weather the renewed wave of scrutiny is by demonstrating its commitment to establishing accountability through program evaluations. Social service agencies will continue to be confronted with an increased demand for their services, while many funders experience significant shortfalls. As social workers progress through the work field, they must therefore be increasingly knowledgeable about seeking and applying for grants, designing logic models and program evaluations, determining outcomes, and successfully implementing evaluations in ways that gather meaningful data.

Additionally, since funders appear to be increasingly requiring collaboration with other agencies, it is imperative that agencies be able to understand what each other is doing, how they're doing it, and why they're doing it. Social workers must be capacitated with knowledge of how to use evaluation to establish organizational accountability as they face the inevitable prisoner's dilemma relating to collaboration: with increasing requests for collaborations, agencies will be faced with the decision to compete as individuals or collaboratives, and may find themselves increasingly relying on each other's knowledge and resources to hold each other accountable.

APPENDIX A

SURVEY REQUEST EMAIL TEMPLATE

Dear AGENCY LEADER,

Nearly a decade ago human service organizations executives from the Dallas area provided information about their organization's use of program evaluation.* Much has changed in that time, as you know, being a leader in the field yourself at AGENCY, but it is not clear exactly how the use of program evaluation has changed in North Texas, specifically when comparing the Dallas and Tarrant county areas.

You can help document the current situation by completing the survey through the link below. The survey should take no more than 15 minutes to complete. As a thank you for your time, if you complete the survey, I will send you a report describing the complete results. This will keep you abreast of how program evaluation is being used and viewed by Metroplex area human service organizations. (The use of SSL technology will allow your information to be submitted electronically and confidentially.) The compiled results may be presented at academic conferences or published, but all individual responses will be kept confidential. The study has been approved by the Institutional Review Board at the University of Texas at Arlington. By completing this survey, you provide your informed consent and grant permission for me to use the results in aggregated form. If you have any questions regarding this study, you may call me at (817)597-1652 or email me at marcela.nava@mavs.uta.edu.

In addition to participating in the research by completing the survey through the link below, you may also volunteer to be interviewed individually (on the phone or in-person). The purpose of this interview would be to gather more detailed information about your organization's use of program evaluation and to round out the quantitative information from the survey. If you would like to volunteer to provide additional information, you can indicate this in the survey.

ONLY YOU have the information that is needed to answer these questions. Please take the few minutes required to complete this survey.

SURVEY LINK: http://tinyurl.com/akhh4w

Marcela Nava, LBSW, UTA Social Work Graduate Student

*The original study was published as: Richard Hoefer, (2000), "Accountability in Action? Program Evaluation in Nonprofit Human Service Agencies," Nonprofit Management and Leadership, 11(2), 167-177. You can read the abstract and download PDF copy of the paper for from http://www3.interscience.wiley.com/journal/104546591/abstract

APPENDIX B

SURVEY

Survey of Social Agency Program Evaluations

1. Informed Consent

Thank you for agreeing to participate in the survey!

* 1. Please enter your contact information below. (If you choose not to be interviewed, this information will only be used to send you a copy of the results at the end of this study.)

Name:
Company:
Address 1:
Address 2:
City/Town:
State:
ZIP/Postal Code:
Email Address:
Phone Number:

2. Interview Consent

If you would like to volunteer to be individually interviewed (in person or on the phone) regarding program evaluation use in your organization, please select the appropriate option below. This study has been approved by the Institutional Review Board at the University of Texas at Arlington. If you have any questions regarding this project, please contact me, Marcela Nava, at (817) 597-1652, or via email at marcela.nava@mavs.uta.edu.

By selecting, "Yes, I agree to be interviewed" below, you consent to be contacted and to have your deidentified comments included in any report of the results. If you do not wish to volunteer to be interviewed, please click, "No, I do not wish to be interviewed".

* 2. Please indicate your willingness to be interviewed below.

- ○ No, I do not wish to be interviewed.
- ○ Yes, I agree to be interviewed.

Please indicate if you have a preference to be interviewed in-person or by phone.

3. PART 1

This set of questions asks about you and your organization.

3. What is your position?

- ○ Executive Director (or similar)
- ○ Other top manager
- ○ Middle Manager
- ○ Other (please specify)

4. How long have you been in your present position?

5. How long have you worked for this organization?

[   ]

6. What are your academic degree(s) and dates?

High School Diploma   [   ]

Associate Degree   [   ]

Bachelor's Degree   [   ]

Master's Degree   [   ]

Doctorate   [   ]

7. What is your age?

[   ]

8. What is your race/ethnicity?

[   ]

9. What is your gender?

◯ Female   ◯ Male

10. What type of human service agency are you?

◯ Nonprofit   ◯ Government   ◯ Forprofit

11. Would you describe your organization as "faith-based"?

◯ No   ◯ Yes

12. When was your organization founded?

[   ]

13. What is your organization's budget in its most recent fiscal year?

[   ]

14. What are the approximate percentages of income from each of the following sources?

Government   [   ]

United Way   [   ]

Foundations   [   ]

Other (specify source and amount)   [   ]

4. PART 2: Agency Use of Program Evaluation

This set of questions asks about your agency's use of program evaluation with its largest human service program during the last two years.

15. About how many non-duplicated clients does your largest human service program have in a typical year?

[   ]

47

16. Has the largest human services program in your agency been formally evaluated at least once during the last two years?

◯ No

◯ Yes

17. If the largest program has NOT been formally evaluated at least once in the last two years, please mark all the reasons why an evaluation has not been conducted from the list below (skip this question if it does not apply).

☐ There is not enough staff time to conduct an evaluation.

☐ Proper knowledge about conducting an evaluation is lacking in the agency.

☐ There is a lack of funds to pay for an evaluation.

☐ The funder of the program does not require an evaluation.

☐ There is no need to conduct an evaluation.

☐ Other (please describe):

[   ]

5. Stop Here if You Answered "NO" to Number 16.

ONLY PROCEED IF YOUR LARGEST PROGRAM HAS BEEN EVALUATED AT LEAST ONCE IN THE LAST TWO YEARS

18. Please mark all the reasons why an evaluation has been conducted from the list below.

☐ To see how the program is doing

☐ To comply with a funder's demands/requirements

☐ To ensure that the program complies with proper procedures

☐ To show funders and potential funders how successful the program is

☐ Other (please describe):

[   ]

19. What is the best characterization of the source of expertise used for designing, conducting and writing the evaluation?

◯ Internal to the agency   ◯ External to the agency   ◯ Both internal and external to the agency

20. Where does funding to conduct the evaluation come from? (CHECK ALL THAT APPLY)

☐ Internal sources

☐ Part of program grant

☐ A separate grant to conduct an evaluation

☐ Evaluation was free (used donated labor, pro bono expertise, etc.)

☐ Other (please describe):

_____

21. Program evaluation encompasses a number of different activities. Please mark the type of activities that your evaluation encompassed (CHECK ALL THAT APPLY)

☐ IMPLEMENTATION MONITORING/FIDELITY ASSURANCE (when program activities are checked to ensure that the program is being delivered to the target population as designed)

☐ PROCESS EVALUATION (when the program activities and procedures are checked to assist in explaining why a program did or did not achieve expected outcomes)

☐ OUTCOME EVALUATION (when the program results are checked to see if program goals are achieved)

☐ OTHER TYPE (please describe):

_____

22. Program evaluations use different research designs. Which have you used to evaluate the largest program in your agency in the last two years? (CHECK ALL THAT APPLY)

☐ POSTTEST ONLY (the clients were measured only once after the program was over)

☐ ONE GROUP PRETEST AND POSTTEST (the clients were measured once both before and after the program was provided to them)

☐ TIME SERIES (the clients were measured more than once before and after the program was provided to them)

☐ COMPARISON GROUP POSTTEST ONLY (clients and another, similar, group were measured after the program was given to the client group)

☐ COMPARISON GROUP PRESTEST AND POSTTEST (clients and another, similar group were measured before and after the program was provided to the client group)

☐ Any type of EXPERIMENTAL EVALUATION (where the people in the study were randomly assigned to either getting the program treatment or a control group)

☐ OTHER DESIGN (please descibe):

_____

23. What type(s) of data sources were used in the evaluation? (CHECK ALL THAT APPLY)

☐ Agency records (client files, organizational reports, etc.)

☐ Non-standardized instruments (surveys or tests made up for the evaluation)

☐ Standardized instruments (scales tested for validity and reliability)

☐ Social indicators (government statistics such as poverty or birth rates)

☐ Other (please descibe):

_____

24. Once an evaluation is conducted, the results can be used in many ways. Check ALL that your agency has used or will use the results for.

☐ To make improvements in the program being evaluated

☐ To verify program goal or outcome achievement

☐ To shift funding away from the program being evaluated

☐ To shift funding to the program being evaluated

☐ To advocate for more money from original funders

☐ To advocate for more money from new funders

☐ Other (please describe):

_____

25. Which stakeholders have formally received information on the results of the program evaluation (via a written or oral report, for example)? (CHECK ALL THAT APPLY)

☐ Managers or administrators other than the program managers

☐ Front-line staff involved in providing the program

☐ The organization's board of directors or similar group

☐ The general public (via posting on website or publication in another form)

☐ Funders of the program

☐ The press or other media

☐ Other (please describe):

_____

26. Who makes the final decision regarding how much of the evaluation's conclusions and recommendations are released and to whom?

○ Executive Director of the organization

○ The organization's Board of Directors or similar group

○ The program evaluator(s)

○ Other (please describe):

```
[                                    ]
```

6.

This is the final section of the survey.

27. Program evaluations can be resisted and supported for many reasons. Please rate how much these statements are generally agreed or disagreed with by the staff in your agency.

	Very much disagree	Fairly much disagree	Slightly disagree	Neither disagree nor agree	Slightly agree	Fairly much agree	Very much agree
No evaluation is needed because we already know that the program is doing a good job.	○	○	○	○	○	○	○
If the results of an evaluation are negative, the evaluation is probably flawed.	○	○	○	○	○	○	○
An evaluation can inform funders of good things the program does.	○	○	○	○	○	○	○
Evaluation results are generally of no use to direct practice staff.	○	○	○	○	○	○	○
An evaluation is too much work.	○	○	○	○	○	○	○
Evaluations merely state the obvious.	○	○	○	○	○	○	○
An evaluation can help improve the program.	○	○	○	○	○	○	○
Evaluations are inherently more trouble than they are worth.	○	○	○	○	○	○	○
The cost of an evaluation is money well-spent.	○	○	○	○	○	○	○
The strengths and weaknesses of the program are already known.	○	○	○	○	○	○	○
Evaluations hurt funding if the results are negative.	○	○	○	○	○	○	○
Evaluations are usually biased or inaccurate.	○	○	○	○	○	○	○
Broadly speaking, evaluations are useful.	○	○	○	○	○	○	○
Evaluations are conducted regularly in this organization.	○	○	○	○	○	○	○

Survey of Social Agency Program Evaluations

7. CONCLUSION

Thank you for your participation in this survey. If you have any questions or comments regarding the survey or the research study, feel free to contact the researcher, Marcela Nava, at (817)597-1652 or at marcela.nava@mavs.uta.edu. A copy of the research results will be sent to you after the study is completed.

APPENDIX C

INTERVIEW PROTOCOL AND QUESTIONS

Thank you for taking the time to talk with me today. As you remember, this interview is a follow-up to the survey regarding program evaluation that you previously completed. Today I'd like to talk with you about the general area of program accountability, with you being an expert on how your organization stays accountable to its various stakeholders. As you know, this research has been approved by the Institutional Review Board at UTA. Your comments are confidential, which means they will only be presented in aggregate form and will not be directly linked to you or your organization. I want to stress that your participation is voluntary and you are free to withdraw your consent to this interview at any time. With this in mind, shall we begin?

1. First, what does the idea of accountability mean to you and your organization?

2. How do you see your organization being accountable to society at large?

3. How do you see your organization as being accountable to its clients?

4. When evaluations have been done in your organization, what is that process like, from the design to the implementation of the evaluation?

5. How are the results of the evaluation disseminated?

6. How do you see the use of program evaluation in providing an accountability mechanism?

7. What are some specific ways that evaluations have been used to improve accountability in your organization?

8. What role does "coercion" from funders play in your decision to conduct evaluations?

9. Do you think that conducting evaluations gives you an edge when seeking funds?

10. Where do you see the field moving, in terms of evaluation of programs?

Those are all the questions I have. Is there any other information relating to these questions that you would like to include in the interview?

Debriefing: What questions do you have for me about the survey, the interview, or any other aspect of this research project?

REFERENCES

Annie E. Casey Foundation. (2007). Kids Count Data Center: The Right Start Online. [Data file].

    Available from Annie E. Casey Foundation Web site, http://www.kidscount.org

Bailey, M. (2005). Think "results," not "evaluation": Before learning the "nuts and bolts" of how

    to do evaluation, nonprofit professionals must make a shift from viewing it as a negative

    to viewing it as a valuable aspect of organizational effectiveness. Public Manager,

    34(1), 8-10.

Benjamin, L.M. (2008). Bearing more risk for results: Performance accountability and nonprofit

    relational work [Electronic version]. Administration & Society, 39(8), 959-983.

Brilliant, E. & Young, D.R. (2004). The changing identity of federated community service

    organizations. Administration in Social Work, 28 (3/4), 23-46.

Brown, L. & Troutt, E. (2007). Reporting does not equal accountability! The importance of

    funding characteristics and process on accountability [Electronic version]. International

    Journal of Public Administration, 30(2), 208-225.

Cutt, J. & Murray, V. (2000). Accountability and effectiveness evaluation in nonprofit

    organizations. New York: Routledge.

Cutt, J. & Ritter, R. (1984). Public nonprofit budgeting: The evolution and application of zero-

    base budgeting. Toronto, Ont.: Institute of Public Administration of Canada.

Dartington, T. (1998). From altruism to action: Primary task and the not-for-profit organization.

    Human Relations, 51, 1477-1493.

Dick, B. (2005). Grounded theory: A thumbnail sketch. [On line] Available at

    http://www.scu.edu.au/ schools/gcm/ar/arp/grounded.html

Editorial: Nonprofit efficiency: United Way seeks cost-effectiveness. (2007, July 6). Newsday,

    (Melville, NY), Retrieved January 7, 2009, from Newspaper Source database.

Fine, A.H., Thayer, C.E., & Colghan, A.T. (2000, Spring). Program evaluation practice in the nonprofit sector. Nonprofit Management and Leadership, 10(3), 331-339. Retrieved January 12, 2008 from Business Source Complete database.

Fitzpatrick, J.L., Sanders, J.R., & Worthen, B.R. (2003). Program evaluation: Alternative approaches and practical guidelines. Boston, MA: Allyn & Bacon.

Frumkin, P. & Kim, M.T. (2002). The effect of government funding on nonprofit administrative efficiency: An empirical test. Retrieved December 10, 2008 from Harvard University, Ash Institute for Democratic Governance and Innovation Web site: http://www.ashinstitute.harvard.edu/Ash/PAQ.pdf

Gibelman, M.G. & Gelman, S.R. (2007, July). Very public scandals: An analysis of how and why nongovernmental organizations get in trouble (A Working Paper). Paper presented at the Fourth International Conference of the International Society for Third-Sector Research, Dublin, Ireland.

Giogia, D. & Dziadosz, G. (2008). Adoption of evidence-based practices in community mental health: A mixed-study of practitioner experience [Electronic version]. Community Mental Health Journal, 44, 347-357.

Glaser, B.G. (1999). Keynote address from the Fourth Annual Qualitative Health Research Conference [Electronic version]. Qualitative Health Research, 9(6), 836-845.

Hall, R. (2008). The evolution of social work practice: Implications for the generalist approach [Electronic version]. International Journal of Social Welfare, 17, 390-395.

Hoefer, R. (2000, Winter). Accountability in action? Program evaluation in nonprofit human service agencies. Nonprofit Management & Leadership, 11(2), 167-177.

Hoefer, R. (2008, November). Using program evaluation for nonprofit accountability: A follow-up study. Paper presented at the 2008 ARNOVA Conference, Philadelphia, PA.

Jurik, N.C. (2004, February). Imagining justice: Challenging the privatization of public life [Electronic version]. Social Problems, 51(1), 1-15.

Kearns, K.P. (1996). Managing for accountability: Preserving the public trust in public and nonprofit organizations. San Francisco, CA: Jossey-Bass.

Leach, M., & Gillham, D. (2008, October). Evaluation of the Evidence-Based practice Attitude and utilization SurvEy for complementary and alternative medicine practitioners [Electronic version]. Journal of Evaluation in Clinical Practice, 14(5), 792-798.

Lindgren, L. (2001, July). The nonprofit sector meets the performance-management movement. Evaluation, 7(3), 285-304.

Lingard, L., Albert, M., & Levinson W. (2008). Grounded theory, mixed methods, and action research. British Medical Journal, 337, 567.

Rosen, A., Proctor, E.E., Morrow-Howell, N., & Staudt, M. (1995, October). Rationales for practice decisions: Variations in knowledge use by decision task and social work service. Research on Social Work Practice, 5(4), 501-523.

Rosen, A., Proctor, E.K. & Staudt, M.M. (1999). Social work research and the quest for effective practice [Electronic version]. Social Work Research, 23, 4-14.

Royse, D., Thyer, B.A., Padgett, D.K. & Logan, T.K. (2006). Program evaluation: An introduction (4th ed.). Belmont, CA: Thomson Brooks/Cole.

Smith, P.C. & Richmond, K.A. (2007, July). Call for greater accountability within the U.S. nonprofit sector. Academy of Accounting and Financial Studies Journal. Retrieved December 29, 2008, from http://findarticles.com/p/articles/mi_hb6182/is_/ai_n29362864

Tassie, B., Murray, V. & Cutt, J. (1998). Evaluating social service agencies: Fuzzy pictures of organizational effectiveness [Electronic version]. International Journal of Voluntary and Nonprofit Organizations, 9(1), 59-79.

Walden, G.I. (2006, December). Who's watching us now? The nonprofit sector and the new government by surveillance. Nonprofit & Voluntary Sector Quarterly, 35(4), 715-720.

Witkin, S.L. & Harrison, W.D. (2001, October). Whose Evidence and for What Purpose? [Electronic version]. Social Work, 46(4), 293-296.

Wholey, J.S., Hatry, H.P., Newcomer, K.E. (2004). Handbook of practical program evaluation.

San Francisco, CA: Jossey-Bass.

BIOGRAPHICAL INFORMATION

Marcela Nava was born and raised in Fort Worth, Texas on October 19, 1985, the daughter of Jorge and Ruth Nava. She grew up in a working class family with parents from Mexico and Honduras, which shaped her interests in social justice and immigration. Marcela graduated with honors from Texas Christian University in May 2007 with a Bachelor of Arts in Social Work. She majored in social work with minors in psychology and criminal justice. In August 2006, she entered the University of Texas at Arlington's School of Social Work. She will graduate with a Master of Science in Social Work with a concentration in Community and Administrative Practice. Marcela is actively involved with social justice activities in her community and plans to pursue a career in public policy.